

10/081,527



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

#4

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2001年 2月28日

出 願 番 号  
Application Number:

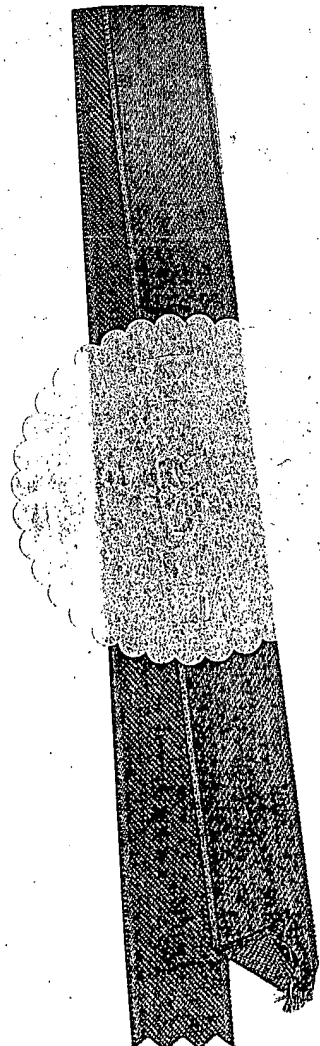
特願2001-053459

[ST.10/C]:

[JP2001-053459]

出 願 人  
Applicant(s):

株式会社日立製作所

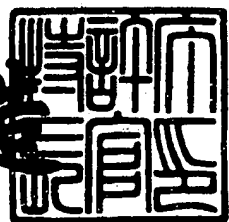


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2002年 1月11日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3114262

【書類名】 特許願

【整理番号】 K01000911A

【提出日】 平成13年 2月28日

【あて先】 特許庁長官

【国際特許分類】 G06F 3/06

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 八木沢 育哉

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 松並 直人

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 味松 康行

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 萬年 暁弘

【発明者】

    【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

    【氏名】 村岡 健司

【特許出願人】

    【識別番号】 000005108

    【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶装置システム及び記憶装置システムにおけるデータの複写方法

【特許請求の範囲】

【請求項 1】

n 個の記憶装置及び m 個の記憶装置に区分された複数の記憶装置を有する記憶装置システムにおいてデータを複写する方法であって、  
前記 n 個の記憶装置から冗長データを作成するために必要な長さのデータを読み出し、  
前記読み出されたデータから冗長データを作成し、  
前記読み出されたデータ及び前記冗長データを前記 m 個の記憶装置に格納することを特徴とするデータ複写方法。

【請求項 2】

前記格納するステップでは、前記読み出されたデータのうちの一部のデータのみが前記 m 個の記憶装置に格納されることを特徴とする請求項 1 記載のデータ複写方法。

【請求項 3】

外部装置と接続され、n 個の記憶装置及び m 個の記憶装置に区分された複数の記憶装置を有する記憶装置システムにおいてデータを複写する方法であって、  
前記外部装置から送信されたデータを前記 n 個の記憶装置及び前記 m 個の記憶装置に二重化して格納し、  
前記外部装置からの指示に基づいてデータの二重化を停止し、  
データの二重化を停止している間に前記外部装置から送信されたデータを、格納場所に関する情報を記録しながら前記 n 個の記憶装置に格納し、  
前記外部装置からの指示に基づいてデータの二重化を再開し、  
前記 n 個の記憶装置から m-1 単位のデータを読み出し、  
前記 m-1 単位のデータから冗長データを作成し、  
前記 m-1 単位のデータのうち、前記記録された情報で指定されるデータ及び前記冗長データを前記 m 個の記憶装置に格納することを特徴とするデータ複写方法

## 【請求項 4】

n 個の記憶装置及び m 個の記憶装置に区分された複数の記憶装置と、  
前記複数の記憶装置を制御する制御部と、  
前記 n 個の記憶装置を制御する n 制御手段と、  
前記 m 個の記憶装置を制御する m 制御手段とを有し、  
前記 n 個の記憶装置から前記 m 個の記憶装置にデータを複写する場合に、  
前記 n 制御手段は、前記 n 個の記憶装置から m - 1 単位のデータを読み出し、  
前記 m 制御手段は、読み出された m - 1 単位のデータに基づいて冗長データを作成し、  
前記読み出された m - 1 単位のデータのうちのいずれか及び前記冗長データを前記 m 個の記憶装置のいずれかに格納することを特徴とする記憶装置システム。

## 【請求項 5】

前記単位は、論理ブロックアドレスを基準とした所定のデータ長であることを特徴とする請求項 4 記載の記憶装置システム。

## 【請求項 6】

外部装置と接続されるインターフェースと、  
前記 n 個の記憶装置及び前記 m 個の記憶装置にデータを二重化して格納する二重化手段と、  
前記外部装置からの指示に従って、前記二重化手段を停止する手段と、  
前記二重化手段が停止している間に、前記インターフェースから入力されるデータが前記 n 個の記憶装置に格納される場所に関する情報を記録する更新管理手段とを有し、  
前記 m 制御手段は、前記更新管理手段に格納された情報に基づいて前記読み出された m - 1 単位のデータのうちのいずれかを前記 m 個の記憶装置のいずれかに格納することを特徴とする請求項 4 記載の記憶装置システム。

## 【請求項 7】

n 個の記憶装置及び m 個の記憶装置に区分された複数の記憶装置と、  
前記複数の記憶装置を制御する制御部と、

前記  $n$  個の記憶装置を制御する  $n$  制御手段と、  
前記  $m$  個の記憶装置を制御する  $m$  制御手段と、  
前記  $n$  個の記憶装置及び前記  $m$  個の記憶装置とで、データを二重化して格納する  
データ二重化手段とを有することを特徴とする記憶装置システム。

【請求項 8】

$n$  と  $m$  は異なる整数であることを特徴とする請求項 4、5、6 及び 7 のうちいずれか一つの記憶装置システム。

【請求項 9】

$n$  個の記憶装置を有する第一の記憶装置システムと、  
前記第一の記憶装置に接続され、 $m$  個の記憶装置を有する第二の記憶装置システムを有し、  
前記第一の記憶装置システムは、  
前記  $n$  個の記憶装置から  $m - 1$  単位のデータを読み出す手段を有し、  
前記第二の記憶装置システムは、  
前記第一の記憶装置システムが読み出した  $m - 1$  単位のデータに基づいて冗長データを作成する作成手段と、  
前記作成手段によって作成された前記冗長データ及び前記読み出された  $m - 1$  単位のデータを前記  $m$  個の記憶装置に格納する手段とを有することを特徴とするシステム。

【請求項 10】

コンピュータと、複数の記憶媒体を有する記憶装置システムとを有する計算機システムであって、  
前記記憶装置システムは、  
ミラー元 LU と、  
ミラー先 LU と、  
前記ミラー元 LU の RAID 制御を行う  $n$  - RAID 制御サブプログラムと、  
前記ミラー先 LU の RAID 制御を行う  $m$  - RAID 制御サブプログラムと、  
前記コンピュータがデータ書き込みを要求したときに前記ミラー元 LU と前記ミラー先 LU に前記データを書き込んで二重化する LU ミラーサブプログラムと、

前記ミラー元LUと前記ミラー先LUの二重化を停止しているときに前記ミラー元LUに対するデータ更新を監視する非ミラー時更新監視サブプログラムと、  
 前記ミラー元LUに対する前記データ更新の更新位置を記録する非ミラー時更新位置管理サブプログラムと、  
 前記記録された更新位置のデータを前記ミラー元LUから前記ミラー先LUにコピーすることで内容を一致させるミラー再同期サブプログラムとを有し、  
 前記ミラー元LU及び前記ミラー先LUは、各々RAID構成が異なることを特徴とする計算機システム。

【請求項 1 1】

前記ミラー元LUは、 $nD + 1P$ で構成され、  
 前記ミラー先LUは、 $mD + 1P$ で構成され、  
 $m$ および $n$ は2以上の整数であって、 $m$ と $n$ は異なる値であることを特徴とする請求項 1 0 記載の計算機システム。

【請求項 1 2】

前記ミラー再同期サブプログラムは、前記記録された更新位置のデータを前記ミラー元LUから前記ミラー先LUにコピーすることで内容を一致させる処理を実行し、  
 前記 $m$ -RAID制御サブプログラムは、前記記録された更新位置のデータを含む前記ミラー先LUのストライプ列のデータが揃うように前記ミラー元LUからデータを読み出す処理を尾Kなうことを特徴とする計算機システム。

【請求項 1 3】

コンピュータと、複数の記憶媒体を持つ第一の記憶装置システム及び第二の記憶装置システムとを有する計算機システムであって、  
 第一の記憶装置システムは、  
 ミラー元LUと、  
 前記ミラー元LUのRAID制御を行う $n$ -RAID制御サブプログラムと、  
 前記コンピュータがデータ書き込みを要求したときに前記ミラー元LUと前記ミラー先LUに前記データを書き込んで二重化するLUミラーサブプログラムと、  
 前記ミラー元LUと前記ミラー先LUの二重化を停止しているときに前記ミラー

元LUに対するデータ更新を監視する非ミラー時更新監視サブプログラムと、  
前記ミラー元LUに対する前記データ更新の更新位置を記録する非ミラー時更新  
位置管理サブプログラムと、  
前記記録された更新位置のデータを前記ミラー元LUから前記ミラー先LUにコ  
ピーすることで内容を一致させるミラー再同期サブプログラムと、  
第二の外部記憶装置とのデータ転送を実施するためのコマンドを発行するコマン  
ド発行サブプログラムとを有し、  
前記第二の記憶装置システムが、  
ミラー先LUと、  
前記ミラー先LUのRAID制御を行うm-RAID制御サブプログラムとを有  
し、  
前記ミラー元LU及びミラー先LUは、RAIDレベルが異なることを特徴とす  
る計算機システム。

【請求項14】

前記ミラー元LUは $nD+1P$ で構成され、  
前記ミラー先LUは $mD+1P$ で構成され、  
 $m$ 及び $n$ は2以上の整数であって、 $m$ と $n$ は異なる値であることを特徴とする請  
求項13記載の計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、コンピュータの外部記憶装置システムにおける記憶領域の管理方法  
、特に、スナップショットについての管理方法に関する。

【0002】

【従来の技術】

ハードディスクなどの記憶装置では、記憶装置の障害、ソフトウェアプログラ  
ムの欠陥、誤操作などによりデータが喪失される場合に備えて、記憶装置に記録  
されたデータを、定期的にテープなどの別の記憶装置にコピーして保存すること  
(バックアップ)が必要となる。その際、バックアップ中にバックアップの対象



となるデータ（オリジナルデータ）が更新され、バックアップされたデータとバックアップされる前のデータとの間に不整合が生じると、バックアップを開始した時点でのデータを再現できない。このため、コピー作業中はデータの整合性を保証する必要がある。

【 0 0 0 3 】

バックアップされるデータの整合性を保証するためには、コンピュータシステムにおいて、バックアッププログラムが動作している間は、バックアッププログラム以外のプログラムを停止すればよい。しかし、高い可用性が要求されるシステムでは、プログラムを長時間停止させることができない。そのため、バックアップ中にプログラムがデータを更新することを妨げず、なおかつバックアップ開始時点での記憶装置に格納されたデータの内容及びデータの配置場所（以下、「記憶イメージ」と称する）をバックアップ用の記憶装置に作成する仕組みが必要となる。ここで、ある時点でのデータの記憶イメージをスナップショットと呼ぶ。また、指定された時点のスナップショットを作成しつつデータの更新が可能な状態を提供する仕組みをスナップショット管理方法と呼ぶ。また、スナップショット管理方法によりスナップショットを作成することをスナップショットの取得と呼ぶ。また、スナップショットを作成した状態をやめることをスナップショットの削除と呼ぶ。

【 0 0 0 4 】

従来のスナップショット管理方法の一つとして、データを二重化する方法がある。

【 0 0 0 5 】

この方法では、スナップショットを取得していない通常の状態において、コンピュータ上のプログラムがすべてのデータを2つの記憶領域に二重化して記憶する。スナップショットを取得する時は、二重化が停止され、1つの領域がオリジナルデータ、もう1つの領域がスナップショットとして提供される。

【 0 0 0 6 】

二重化が停止されている間にオリジナルデータの更新が発生した場合、更新されたデータの位置が記録される。スナップショット削除時には、データの二重化

が再開される。また、更新されたデータ位置の記録をもとに、オリジナルデータの記憶領域からもう一方の記憶領域に更新されたデータがコピーされる。このコピー動作を、ミラー再同期化と呼ぶ。コンピュータ上のプログラムでデータを二重化する方法は、米国特許 5, 0 5 1, 8 8 7 号に開示されている。

#### 【 0 0 0 7 】

上記 2 つの記憶領域は、一般にディスクアレイによって構築されることが多い。ディスクアレイは、複数のディスク装置をアレイ状に配置した構成を有する。ディスクアレイを含んだ記憶装置システムを、以下、ディスクアレイ装置と呼ぶ。ディスクアレイ装置は、ホスト装置（ホスト）からのリード要求及びライト要求をディスク装置の並列動作によって高速に処理するとともに、データに冗長データを付加することによって信頼性を向上させた記憶装置である。ディスクアレイは、RAID (Redundant Arrays of Inexpensive Disks) とも呼ばれる。ディスクアレイは、冗長データの種類と構成により一般的に 5 つのレベルに分類される。

#### 【 0 0 0 8 】

スナップショットが使用されている間、オリジナルデータが格納されている記憶領域に対するアクセスの速度が低下しないように、オリジナルデータの記憶領域と、スナップショットの記憶領域は、それぞれ物理的に異なるディスクアレイで構成される。また、ディスクアレイにおけるディスク装置の構成 ( $nD + 1P$  :  $D$  はデータが格納されるディスク装置を示し、 $P$  は冗長データが格納されるディスク装置を示す。 $nD + 1P$  は、 $n$  台のデータを記録するディスク装置及び 1 台の冗長データが格納されるディスク装置の構成を示す) で考えると、オリジナルデータの記憶領域が配置されるディスクアレイと、スナップショットに使用される記憶領域が配置されるディスクアレイは、同じディスク装置の数、すなわち、同じ  $n$  を持つ構成とされ、個々のディスク装置の記憶容量が同一のものが使用される。

#### 【 0 0 0 9 】

##### 【発明が解決しようとする課題】

オリジナルデータが格納される記憶領域と、スナップショットに使用される記

憶領域とでは、性能、信頼性、コスト等の観点から、異なる特徴が要求される。例えば、オリジナルデータが格納される記憶領域では信頼性が重視され、スナップショットに使用される記憶領域には低いビットコストが要求される。また、バックアップを高速化するため、スナップショットに使用される記憶領域に高性能が要求される場合もある。

#### 【0010】

つまり、性能、記録密度、容量、信頼性等が異なるディスク装置で各ディスクアレイが構成される必要がある。この場合、ディスクアレイの構成 ( $nD+1P$ ) の仕方によっても、要求される特徴を実現することができるので、オリジナルデータの記憶領域と、スナップショットに使用される記憶領域が配置される各々のディスクアレイを互いに異なる  $n$  を持つ  $nD+1P$  で構成する必要がある。しかし、従来技術においては、互いに柔軟な構成が採用されるディスクアレイの組み合わせは実現することが出来なかった。

#### 【0011】

本発明の目的は、バックアップに使用するスナップショットを取得するために二重化して運用されているディスクアレイにおいて、オリジナルデータが格納される記憶領域と、スナップショットに使用される記憶領域を有するそれぞれのディスクアレイを互いに異なる  $n$  を持つ  $nD+1P$  のディスク構成とし、各ディスクアレイが互いに柔軟な構成を採れる方法を提供することである。

#### 【0012】

##### 【課題を解決するための手段】

前記目的を達成するために、本発明は、スナップショット取得のために二重化運用しているディスクアレイにおいて、 $nD+1P$  で構成する複数の記憶媒体上の記憶領域であるミラー元LUと、 $mD+1P$  で構成する複数の記憶媒体上の記憶領域であるミラー先LUと、 $nD+1P$  のRAID制御を行う  $n$ -RAID制御サブプログラムと、 $mD+1P$  のRAID制御を行う  $m$ -RAID制御サブプログラムと、コンピュータがデータ書き込みを要求したときにミラー元LUとミラー先LUに書き込んで二重化するLUミラーサブプログラムとを設ける。なお、 $m$ および $n$ は2以上の整数であって、 $m$ と $n$ は異なる値である。

【0013】

## 【発明の実施の形態】

図1は、本発明が適用されたシステムの第1の実施形態の構成を示す図である。

【0014】

本システムは、コンピュータ100及びディスクアレイ装置200を有する。

【0015】

コンピュータ100は、CPU110、メモリ120、及びSCSIインタフェース140を有する。SCSIインタフェース140は、SCSIバス300を介してディスクアレイ装置200と接続されている。メモリ120には、データベースプログラム126及びバックアッププログラム127が格納されている。両プログラムは、CPU110によって実行される。

【0016】

本実施形態では、スナップショットの利用例として、バックアップをとりあげるが、OLAP (OnLine Analytical Processing) やシステムテスト等の他の目的においても本発明は利用可能である。

【0017】

データベースプログラム126は、後述するLU261にアクセスする機能、及びLU261へのデータ更新を制御し、LU261内のデータの整合性を保証するバックアップモードを実行する機能を有する。データベースプログラム126は、ユーザまたはバックアッププログラム127からの指示によりバックアップモードへ遷移する。バックアッププログラム127は、ユーザからの指示によりスナップショットを保存したLU262からテープ等（図示せず）にデータをバックアップするためにデータを読み出す機能、ディスクアレイ装置200にSCSIのMode Select コマンドを発行する機能、及びデータベースプログラム126にバックアップモードの有効化、無効化を指示する機能を有する。

【0018】

ディスクアレイ装置200は、CPU210、メモリ220、SCSIインタフェース240、ディスクコントローラ250、及びディスク装置群251、2

52を有する。メモリ220内にはスナップショット管理プログラム221が格納されている。本プログラムは、CPU210によって実行される。

#### 【0019】

ディスク装置群251は、ディスク装置271～275を有するディスクアレイである。ディスク装置群252は、ディスク装置281～287を有するディスクアレイである。各ディスク装置群は、パリティ付きのストライピングされたRAID5のディスク構成をとる。尚、本実施形態では、ディスク装置群251のディスク装置の数を5、ディスク装置群252のディスク装置の数を7としているが、各ディスク装置群のディスク装置の数が3つ以上で、かつ、ディスク装置群251とディスク装置群252とが異なるディスク装置の数で構成されていれば、どのような構成であっても良い。以下、ディスク装置群251、252のディスク装置の数をそれぞれ $nD+1P$ 、 $mD+1P$ と表す。本実施形態においては、 $n=4$ 、 $m=6$ である。本発明において、 $n$ 、 $m$ は、2以上の整数であり、 $n$ と $m$ は異なる値とする。

#### 【0020】

ディスク装置群251、252は、複数のディスク装置が有する記憶領域を一つの論理的な記憶領域としてコンピュータ100に提供する。具体的には、SCSI規格に対応した論理ユニット(Logical Unit: LU)がコンピュータ100に提供される。ディスク装置群251、252が提供するLUを、LU261、262とする。ディスク装置群251、252内のLUは複数であってもよい。

#### 【0021】

本実施形態では、ディスクアレイ220内のスナップショット管理プログラム221が、LU261をオリジナルデータを持ったミラー元LUとし、LU262をオリジナルデータのミラーであるミラー先LUとして管理する。LU262が、スナップショットとして使用されるLUである。

#### 【0022】

ディスクアレイ装置200のスナップショット管理プログラム221は、コンピュータ100からの要求に応じてディスクコントローラ250にディスクアク

セスを指示するディスクアクセスサブプログラム230、LU261に対するデータの更新を二重化してLU262のデータも更新し、2つのLUに同じユーザデータを書き込むLUミラーサブプログラム231を有する。

#### 【0023】

また、ディスクアレイ装置200のスナップショット管理プログラム221は、 $n$ -RAID制御サブプログラム236及び $m$ -RAID制御サブプログラム237を有する。 $n$ -RAID制御サブプログラム236は、ディスク装置271~275を $nD+1P$ のRAID5として制御する。 $m$ -RAID制御サブプログラム237は、ディスク装置281~287を $mD+1P$ のRAID5として制御する。

#### 【0024】

さらに、スナップショット管理プログラム221は、ディスクアレイ装置200が二重化の処理を停止している時（非ミラー時）に、ミラー元LU261に対するデータ更新の有無を検出する非ミラー時更新監視サブプログラム234、データの更新位置を後述する更新位置管理表222に記録する非ミラー時更新位置管理サブプログラム235、及びミラー再同期化を行う際にミラー元LUの更新部分をミラー先LUにコピーするミラー再同期サブプログラム232とを有する。

#### 【0025】

図2は、更新位置管理表222の構成を示す図である。更新位置管理表222は、ミラー元LU内のLBAセット番号、及びLBAセット番号に対応する更新ビットを格納する領域を有する。LBAセットは、1個以上のLBA（Logical Block Address）を単位としてLUを先頭から分割していったときの個々の部分領域である。LBAセット番号は、LBAセットに振られた通し番号である。更新ビットは、非ミラー時にミラー元LUのLBAセットが更新されたかどうかを示し、「更新」、「非更新」に応じて各々1、0が指定される。更新ビットの初期設定値は0である。図2の例では、LBAセット番号1の領域のみが非ミラー時に更新されている状態を示す。

#### 【0026】

図3は、スナップショット取得／削除時におけるバックアッププログラム127とスナップショット管理プログラム221の動作を示すフローチャートである。ここでは、ミラー元LUであるLU261のミラー先LUであるLU262を、LU261のスナップショットとして提供する場合を説明する。

【0027】

コンピュータ100のバックアッププログラム127がデータベースプログラム126に指示を与え、バックアップモードを有効化してスナップショットを取得するデータの整合性を保証する(2000)。次に、バックアッププログラム127は、ディスクアレイ装置200にスナップショットを取得するためのModeSelectコマンドを発行する(2001)。ディスクアレイ装置200のスナップショット管理プログラム221は、ModeSelectコマンドを受信すると(3000)、非ミラー時更新監視サブプログラム234と非ミラー時更新位置管理サブプログラム235を有効化し、LU261の更新位置記録を開始する(3001)。以降、LU261のデータが更新されると、更新位置管理表222における更新されたLBAを含むLBAセットの更新ビットが1に設定される。

【0028】

次に、スナップショット管理プログラム221は、LUミラーサブプログラム231を無効化し、LU261とLU262のデータの二重化の処理の停止を指示する。これにより、ミラー元LUであるLU261に対するデータの更新がミラー先LUであるLU262に反映されなくなる(3002)。次に、スナップショット管理プログラム221は、ModeSelectコマンドの終了ステータスをコンピュータ100のバックアッププログラム127に送信する(3003)。バックアッププログラム127は、ModeSelectコマンドの終了ステータスを受信すると(2002)、データベースプログラム126に指示を与え、バックアップモードを無効化する(2003)。

【0029】

次に、バックアッププログラム127は、ディスクアレイ装置200に対し、スナップショット削除を指示するModeSelectコマンドを発行する(2

004)。ディスクアレイ装置200のスナップショット管理プログラム221は、Mode Select コマンドを受信すると(3005)、LUミラーサブプログラム231を有効化し、LU261とLU262の二重化を再開する(3006)。これにより、LU261に対する更新がLU262にも反映される。

#### 【0030】

次に、スナップショット管理プログラム221は、非ミラー時更新監視サブプログラム234と非ミラー時更新位置管理サブプログラム235を無効化し、LU261の更新位置記録を停止する(3007)。以降、更新位置管理表222の更新ビットを非ミラー時更新位置管理サブプログラム235が変更しなくなる。

#### 【0031】

次に、スナップショット管理プログラム221は、ミラー再同期サブプログラム232を有効化する。ミラー再同期サブプログラム232は、更新位置管理表222を参照して、LU261とLU262で内容が一致しない部分をLU261からLU262にコピーすることでミラー再同期化を行う(3008)。再同期が終了したら、スナップショット管理プログラム221は、ミラー再同期サブプログラム232を無効化し(3009)、Mode Select コマンドの終了ステータスをコンピュータ100のバックアッププログラム127に送信する(3010)。Mode Select コマンドの終了ステータスを受信したバックアッププログラム127は、動作を終了する(2005)。

#### 【0032】

コンピュータ100のバックアッププログラム127は、ステップ2003～ステップ2004の間で、スナップショットを取得したLU262の読み出しを行い、他の記憶装置(図示せず)にそのデータをバックアップする。

#### 【0033】

ステップ3008で、LU261からLU262へのデータコピーを行うミラー再同期化の動作について説明する。

#### 【0034】

図5は、LU261とLU262のディスクアレイ構成が、 $n < m$ の場合のL



U 2 6 1、L U 2 6 2、及び更新位置管理表 2 2 2 の関係を示す図である。

【 0 0 3 5 】

L U 2 6 1 は、データの L B A セットが 4 個に対して、パリティを 1 個割り振る構成である。例えば、L B A セット 0 ～ 3 に対してパリティ P n 0 が割り振られている。パリティ P n 0 は、L B A セット 0 ～ 3 の排他的論理和を取ったものである。同様に、L B A セット 4 ～ 7 に対してパリティ P n 1 が、L B A セット 8 ～ 1 1 に対してパリティ P n 2 が割り振られる。ここで、複数の L B A セット及びその L B A セットの組に対応するパリティを合わせてストライプ列と呼ぶ。例えば、L B A セット 0 ～ 3 とパリティ P n 0 でストライプ列が構成される。尚、R A I D 5 では、パリティが各ディスク装置にローテーションしながら分散して配置されるが、本実施形態ではローテーションする単位が大きいと想定し図中には分散配置の状況が表れないものとする。

【 0 0 3 6 】

L U 2 6 2 は、データの L B A セットが 6 個に対して、パリティを 1 個割り振る構成となっている。例えば、L B A セット 0 ～ 5 に対してパリティ P m 0 が割り振られている。

【 0 0 3 7 】

尚、L U 2 6 1 及び L U 2 6 2 の L B A セット番号が同じ記憶領域同士でデータの二重化が実施される。

【 0 0 3 8 】

図 5 では、更新位置管理表 2 2 2 と L U のストライプ列との関係を示すために、更新位置管理表 2 2 2 の左側に L U 2 6 1 と L U 2 6 2 のストライプ区切りを併記した。なお、n と m は値が異なることから、L U 2 6 1 と L U 2 6 2 のストライプ区切りにずれが生ずる。

【 0 0 3 9 】

ミラー再同期化時には、L U 2 6 1 と L U 2 6 2 の間で内容が一致しない L B A セットが、L U 2 6 1 から L U 2 6 2 にコピーされる。このとき、L U 2 6 1 に対してはデータの読み出しの処理が行われ、L U 2 6 2 に対してはデータの書き込みの処理が行われる。

## 【0040】

RAID5の構成であるディスクアレイにデータの書き込みを行う場合、パリティの整合性をとるために、ストライプ列内の更新されないLBAセットのデータを読み出し、新しいデータを含め排他的論理和演算をして新しいパリティを生成した後、新しいパリティ及び新しいデータをディスク装置に書き込む動作が行われる。このように、ディスクアレイへのデータ書き込みにおいて、データを書き込むディスク装置以外にディスクアクセス回数が増える現象をライトペナルティと呼ぶ。

## 【0041】

なお、ストライプ列のすべてのLBAセットのデータが更新される場合、新しいパリティは、書き込まれるLBAセットのみで生成される。つまり、パリティを作成するためにストライプ列内の更新されないLBAセットのデータを読み出す必要がない。したがって、パリティの書き込みが増えるだけなので、最小限のライトペナルティで済ますことができる。

## 【0042】

ここで、更新位置管理表222のLBAセット番号7～10の更新ビットに1が設定されており、ミラー再同期化にて該当するLBAセット7～10のデータをLU261からLU262にコピーする場合を想定する。

## 【0043】

通常にコピーする場合は、LU261の太枠で囲われたLBAセット7～10のデータを読み出し、LU262の太枠で囲ったLBAセット7～10に書き込む。それに加えて、LU262のストライプ列のパリティの整合性をとるため、LU262の斜線で示したLBAセット6とLBAセット11を読み出し、LU261のLBAセット7～10と排他的論理和を演算して生成したパリティをディスク装置287のPm1に書き込む。したがって、パリティの書き込み以外に、LU262のLBAセット6及びLBAセット11の読み出しのためのリードアクセスを行わなければならない。

## 【0044】

本発明では、ライトペナルティを軽減するため、n-RAID制御サブプログ

ラムが、LU262ではなく、LU261から斜線で示したLBAセット6及びLBAセット11のデータを読み出し、LU261のストライプ列のパリティ生成のためのデータを準備する。したがって、LU262の斜線で示したLBAセット6及びLBAセット11のデータを読み出す必要はない。また、LU261のLBAセット7～10のデータを読み出す際、ディスク装置271～274へのディスクアクセスは発生するが、その際、LBAセット10を読み出すときにLBAセット6も連続して1つのディスクコマンドで読み出す。同様に、LBAセット7を読み出すときに、LBAセット11も連続して1つのディスクコマンドで読み出す。ディスクアクセスは、ディスク装置内部の読み書きヘッドの部分のシーク時間とディスク円盤の回転待ち時間が主要な処理時間である。したがって、1つのディスクコマンドでデータを読み出すことでディスク円盤の回転待ち時間を省略することができ、LBAセット6、11へのディスクアクセス時間を削減できる。

## 【0045】

つまり、本発明では、ミラー再同期化のためにコピーが必要なLBAセットがあった場合、n-RAID制御サブプログラム236が、コピーが必要なLBAセットを含むミラー先LUのストライプ列の中のすべてのLBAセットがそろるようにミラー元LUからデータの読み出しを行う。

## 【0046】

ミラー元LUからの読み出しが必要なLBAセットの算出は以下のように行う。本実施形態においては、ミラー先LUであるLU262が、 $m=6$ である。したがって、ミラー先LUの1つのストライプ列は、 $p$ を0以上の整数として、 $6 \times p \sim 6 \times p + 5$ のLBAセット番号で指定されるLBAブロックで構成される。ミラー再同期化のためのコピーが必要なLBAセットが含まれるストライプ列内のLBAセット番号は、ミラー再同期化のためのコピーが必要なLBAセット番号を $m=6$ で除算した整数部分として $p$ を求めることで特定される。

## 【0047】

例えば、LBAセット番号7に対応する更新ビットが1である場合、7を6で除算した整数部分、すなわち $p=1$ となり、読み出しが必要なLBAセット番号

は、 $6 \times p \sim 6 \times p + 5$ 、すなわち、LBAセット番号6～11となる。 $n$ -RAID制御サブプログラムは、この $p$ の値から特定されたLBAセット番号に基づいて読み出すべきLBAセットの位置を決定し、データを読み出す。

## 【0048】

ミラー先LUにおいては、 $m$ -RAID制御サブプログラムが、コピーが必要なLBAセットの書き込み、及び生成したパリティの書き込みを実施する。

## 【0049】

なお、コピーが必要なLBAセットの検出は、ミラー再同期サブプログラム232が更新位置管理表222を参照することにより行う。ミラー再同期サブプログラム232は、ミラー先LUのストライプ列内でコピーが必要なLBAセットがあった場合、同じストライプ列内にコピーが必要な他のLBAセットがあるかどうかを検索し、他のLBAセットもコピーするように動作する。具体的には、ミラー再同期サブプログラム232は、更新位置管理表222の更新ビットが1となっているLBAブロックを検索する。ビットが1であるLBAセットを発見したら、ミラー再同期サブプログラム232は、ミラー先LUにおいて、発見されたLBAセットを含むストライプ列内の他のLBAセットに対応する更新ビットを参照し、1であるLBAセットの有無を確認する。ミラー再同期サブプログラム232は、同じストライプ列内において更新ビットが1であるLBAブロックを確認した後、該当するLBAブロックについてコピーをするよう、 $m$ -RAID制御サブプログラムに指示を送る。また、 $n$ -RAID制御サブプログラムに $p$ の値を通知し、 $p$ に対応するLBAブロックを読み出すように指示を送る。

## 【0050】

尚、ミラー先LUの1つのストライプ列ごとに、ミラー再同期サブプログラム232が更新位置管理表222を検索するとして説明したが、コピーが必要なLBAセットが、ミラー先LUにおいて隣接するストライプ列にまたがっている場合は、隣接する複数のストライプ列内においてもコピーが必要な他のLBAセットがあるかどうかを検索し、他のLBAセットもコピーするように動作するようにしてもよい。

## 【0051】

また、ディスクアクセス回数を減らすことができない場合もあるので、通常のコピー、すなわち、更新ビットが1となっているLBAブロックを検出したら、すぐにn-R A I D制御サブプログラム及びm-R A I D制御サブプログラムに当該LBブロックの読み出し及び書き込みを行うよう、LU261の読み出し位置とLU262の書き込み位置を算出してもよい。

#### 【0052】

図6は、LU261とLU262のディスクアレイ構成が、 $n > m$ の場合のLU261、LU262、及び更新位置管理表222を示す図である。ここでは、 $n = 4$ 、 $m = 3$ の場合で説明する。

#### 【0053】

LU261は、図5と同様である。

#### 【0054】

LU262は、データのLBAセットが3個に対して、パリティを1個割り振る構成となっている。

#### 【0055】

図5と同様に、更新位置管理表222の左側に、LU261とLU262のストライプ区切りを併記する。

#### 【0056】

例えば、更新位置管理表222のLBAセット番号4～7の更新ビットが1に設定されており、ミラー再同期化処理にて、該当するLBAセット4～7のデータがLU261からLU262にコピーされる場合を説明する。

#### 【0057】

通常にコピーする場合は、LU261の太枠で囲ったLBAセット4～7を読み出し、LU262の太枠で囲ったLBAセット4～7に書き込む。それとともに、LU262のストライプ列のパリティ整合性をとるためにLU262の斜線で示したLBAセット3とLBAセット8を読み出し、LU261のLBAセット4～5とLBA262から読み出したLBAセット3で排他的論理和演算をし、生成したパリティをディスク装置284のPm1に書き込む。また、LU261のLBAセット6～7とLBA262から読み出したLBAセット8で排他的

論理和演算をして生成したパリティをディスク装置 2 8 4 の P m 2 に書き込む。  
したがって、パリティの書き込み以外に、L U 2 6 2 の L B A セット 3、8 の読み出しのためにディスク装置 2 8 1、2 8 3 へのリードアクセスが発生する。

## 【 0 0 5 8 】

これに対し、本発明ではライトペナルティを軽減するため、n - R A I D 制御サブプログラム 2 3 6 が、L U 2 6 2 ではなく、L U 2 6 1 から斜線で示した L B A セット 3 と L B A セット 8 を読み出し、L U 2 6 1 のストライプ列のパリティ生成のためのデータを準備する。したがって、L U 2 6 2 の斜線で示した L B A セット 3 と L B A セット 8 を読み出す必要がなくなる。また、コピーのための L U 2 6 1 の L B A セット 4 ~ 7 の読み出しの際、ディスク装置 2 7 1 ~ 2 7 4 へのディスクアクセスは発生するが、その際、L B A セット 7 を読み出すときに L B A セット 3 も連続して 1 つのディスクコマンドで読み出す。同様に、L B A セット 4 を読み出すときに L B A セット 8 も連続して 1 つのディスクコマンドで読み出す。1 つのディスクコマンドでデータを読み出すことで、L B A セット 3、8 へのディスクアクセス時間を削減できる。

## 【 0 0 5 9 】

つまり、本発明ではミラー再同期化のためのコピーが必要な L B A セットがあった場合、n - R A I D 制御サブプログラム 2 3 6 が、L B A セットを含むミラー先 L U のストライプ列内のすべての L B A セットがそろうように、ミラー元 L U から代わりとなる L B A セットを読み出す。この際、再同期化のためにコピーが必要となる L B A セットが格納されているミラー元 L U のディスク装置内に、ミラー先 L U の L B A セットの代わりとなる L B A セットがあれば、図 5 と同様に、読み出しにかかる時間が短縮される。ミラー先 L U に対しては、m - R A I D 制御サブプログラム 2 3 7 が、コピーが必要な L B A セットの書き込み、および、生成したパリティの書き込みを実施する。

## 【 0 0 6 0 】

なお、コピーが必要な L B A セットの検出は、図 5 の場合と同様に、ミラー再同期サブプログラム 2 3 2 が更新位置管理表 2 2 2 を参照することにより行う。

## 【 0 0 6 1 】

また、ディスクアクセス回数を減らすことができない場合もあるので、通常のコピーを行うようにLU261の読み出し位置とLU262の書き込み位置を算出してもよい。

#### 【0062】

図4は、ミラー再同期化時におけるミラー再同期サブプログラム232の動作を示すフローチャートである。

#### 【0063】

まず、ミラー再同期サブプログラム232が、更新位置管理表222の更新ビットに更新記録として1があるかどうかを調べる(1001)。もし、更新記録である1がなければミラー再同期化が完了したので処理を終了する(1002)。

#### 【0064】

更新ビットに1がある場合、ミラー再同期サブプログラム232は、図5及び図6で説明したように、LU261の読み出し位置とLU262の書き込み位置を算出する(1003)。次に、該当する読み出し／書き込み位置の更新を抑止し(1004)、算出した読み出し／書き込み位置に基づき、図5及び図6で説明した方法にて、n-RAID制御サブプログラム236及びm-RAID制御サブプログラムが、該当するデータをミラー元LUであるLU261からミラー先LUであるLU262にコピーする(1005)。

#### 【0065】

次に、ミラー再同期サブプログラム232は、該当する読み出し／書き込み位置の更新抑止を解除し(1006)、更新位置管理表222にあるコピーによってミラー化した位置の更新ビットに0を設定して更新記録を削除し(1007)、ステップ1001に戻る。

#### 【0066】

データ読み出しと書き込み、および、スナップショット読み出しの際のコンピュータ100の動作を説明する。

#### 【0067】

まず、コンピュータ100が、ディスクアレイ装置200にあるLU261の

データにアクセスする場合のデータベースプログラム126の動作を説明する。  
データベースプログラム126は、スナップショット取得の有無に関係なく、同じ動作を行う。

## 【0068】

データベースプログラム126がLU261のデータを読み出す場合、データベースプログラム126はディスクアレイ装置200に対し、LU261のデータを読み出すREADコマンドを発行する。データベースプログラム126は、ディスクアレイ装置200からデータとステータスを受信し動作を終了する。また、データベースプログラム126がLU261にデータを書きこむ場合、データベースプログラム126はディスクアレイ装置200に対し、LU261にデータを書きこむWRITEコマンドを発行し、データを送信する。データベースプログラム126は、ディスクアレイ装置200からステータスを受信し動作を終了する。

## 【0069】

次に、コンピュータ100が、ディスクアレイ装置200にあるLU261のスナップショットを読み出す場合のバックアッププログラム127の動作を説明する。

## 【0070】

バックアッププログラム127がLU261のスナップショットを読み出す場合、バックアッププログラム127はディスクアレイ装置200に対し、LU261のミラー先LUであるLU262のデータを読み出すREADコマンドを発行する。バックアッププログラム127は、ディスクアレイ装置200からデータとステータスを受信し動作を終了する。

## 【0071】

データ読み出しと書きこみ、および、スナップショット読み出しの際のディスクアレイ装置200の動作を説明する。

## 【0072】

まず、コンピュータ100が、ディスクアレイ装置200にあるLU261のデータにアクセスする場合のスナップショット管理プログラム221の動作を説



明する。

【0073】

コンピュータ100がLU261のデータを読み出す場合、スナップショット管理プログラム221がLU261に対するREADコマンドを受信する。次に、スナップショット管理プログラム221は、LUミラーサブプログラム231が有効で、かつ、ミラー再同期サブプログラム232による更新部分のコピーが終了していれば、LU261、又はミラー先LUであるLU262からデータを読み出す。そうでない場合、LU261からデータを読み出す。データを読み出したら、スナップショット管理プログラム221は、読み出したデータとステータスをコンピュータ100に送信する。LU261とミラー先LUであるLU262の内容が一致している場合は、両者のいずれかからデータを読み出すことにより負荷が分散される。

【0074】

コンピュータ100がLU261にデータを書き込む場合、スナップショット管理プログラム221がLU261に対するWRITEコマンド及びデータを受信する。次に、スナップショット管理プログラム221は、LUミラーサブプログラム231が有効であればLU261とミラー先LUであるLU262にデータを書きこみ、無効であればLU261にデータを書きこむ。次に、スナップショット管理プログラム221は、非ミラー時更新監視サブプログラム234及び非ミラー時更新位置管理サブプログラム235が有効であれば、LU261の更新位置管理表222に対して更新したデータを含むLBAセットの更新ビットを1に設定し、無効であれば何もしない。最後に、ステータスをコンピュータ100に送信する。

【0075】

次に、コンピュータ100が、ディスクアレイ装置200にあるLU261のスナップショットを読み出す場合のスナップショット管理プログラム221の動作を説明する。

【0076】

コンピュータ100がLU261のスナップショットを読み出す場合、スナッ

プショット管理プログラム221がLU261のミラー先LUであるLU262に対するREADコマンドを受信する。次に、スナップショット管理プログラム221は、ミラー先LUであるLU262からデータを読み出す。データを読み出したら、スナップショット管理プログラム221は、読み出したデータとステータスをコンピュータ100に送信する。

## 【0077】

なお、ミラー再同期サブプログラム232によって更新部分がコピーされている間は、コピー処理とコンピュータ100によるLU261へのデータアクセス処理が同じLU261に集中するため、データアクセス性能が低下する。

## 【0078】

本実施形態によれば、スナップショット取得のために二重化運用しているディスクアレイ装置において、 $n$ と $m$ を異なる2以上の整数とした場合、オリジナルデータの記憶領域を有するディスクアレイと、スナップショットとして提供する記憶領域を有するディスクアレイをそれぞれ $nD+1P$ と $mD+1P$ のディスク構成で組むことが可能となり、各ディスクアレイが互いに柔軟な構成を採ることができるという効果がある。

## 【0079】

したがって、性能、信頼性、コスト等の観点から見たオリジナルデータの記憶領域と、スナップショットとして提供する記憶領域への要求に対応できる。例えば、同じディスク装置を用いて、オリジナルデータの記憶領域を $4D+1P$ 、スナップショットとして提供する記憶領域を $8D+1P$ で構成する場合を考える。信頼性に関しては、冗長なディスク装置数の割合が大きい $4D+1P$ の方が優位であり、ビットコストに関しては、冗長なディスク装置数の割合が小さい $8D+1P$ の方が優位である。ディスクアレイとしての連続したデータの読み出し性能は、ディスク装置が性能のボトルネックとなる場合を想定すると、 $8D+1P$ の方が2倍の性能となり優位である。このように、 $nD+1P$ の異なるディスクアレイでミラー化を図ることで、オリジナルデータの記憶領域と、スナップショットとして提供する記憶領域への様々な要求に対応できるという効果がある。

## 【0080】

また、データの二重化によるスナップショット管理方法では、ミラー再同期化時において、更新／参照をする通常アクセスと更新データのコピーアクセスがオリジナルデータの記憶領域に集中し、通常アクセスの性能が低下するという問題があるが、本実施形態によれば、ミラー先LUのストライプ列に合わせて、ミラー元LUを読み出すことで、性能が低下するミラー再同期化時間を短縮することができる。

#### 【0081】

なお、本発明ではコンピュータ100及びディスクアレイ装置200を接続するインターフェースをSCSIバス300としたが、ファイバチャネル等の他のインターフェースであってもよい。

#### 【0082】

図7は、本発明が適用されたシステムの第2の実施形態の構成を示す図である。本実施形態は、第一の実施形態におけるディスク装置群が別々のディスクアレイ装置で構成されているシステムにおいて本発明を適用したものである。本実施形態では、別々のディスクアレイ装置を用いることから、スナップショット利用の目的だけでなく、遠隔地へのデータのミラーリング目的に本発明を適用することも可能である。以下、第一の実施形態と異なる部分について説明する。

#### 【0083】

本実施形態のシステムは、コンピュータ100、ディスクアレイ装置201、及びディスクアレイ装置202を有する。複数のディスクアレイ装置を有する点が第一の実施形態とは異なる。

#### 【0084】

コンピュータ100は、FCインターフェース190を有し、FCインターフェース190を介して他の装置と接続される点が第1の実施形態と異なる。

#### 【0085】

ディスクアレイ装置201は、ディスク装置群252及びm-R A I D制御サブプログラム237がディスクアレイ装置200から取り除かれ、スナップショット管理プログラム221内に他のディスクアレイ装置等とのデータ転送を実施するためにコマンドを発行するコマンド発行サブプログラム238が設けられて

いる点が第一の実施形態のディスクアレイ装置200とは異なる。また、ディスクアレイ装置202は、SCSIインタフェース240の代わりにFCインタフェース290を有し、FCインタフェース290を介して他の装置と接続される。

【0086】

ディスクアレイ装置202は、ディスクアレイ装置200からディスク装置群251及びスナップショット管理プログラム221が取り除かれ、ディスクアレイプログラム223がメモリ220に設けられた点が第一の実施形態のディスクアレイ装置200とは異なる。また、ディスクアレイ装置202は、SCSIインタフェース240の代わりにFCインタフェース290を有し、FCインタフェース290を介して他の装置と接続される。

【0087】

ディスクアレイプログラム223は、第1の実施形態の図1と同じディスクアクセスサブプログラム230及びm-Raid制御サブプログラム237を有する。

【0088】

ディスクアレイ装置201のコマンド発行サブプログラム238は、ファイバチャネルスイッチ410経由でディスクアレイ装置202内のLU262にアクセスするためのコマンドをディスクアレイ装置202に発行する。したがって、ディスクアレイ装置201は、ファイバチャネルスイッチ410経由で接続されたディスクアレイ装置202にあるLU262を第1の実施形態と同様にミラー先LUとするように動作する。他のプログラムも同様に、ファイバチャネルスイッチ410経由でLU261とLU262でのデータのコピー等を実行する。

【0089】

また、コンピュータ100が、ディスクアレイ装置201にあるLU261のデータにアクセスする場合のスナップショット管理プログラム221の動作は、LU261とミラー先LUであるLU262の内容が一致している場合においても、ミラー元LU261のみ読み出すようにしてもよい。

【0090】

本実施形態によれば、二重化運用されている複数のディスクアレイ装置において、 $n$ と $m$ を異なる2以上の整数とした場合、オリジナルデータの記憶領域を有するディスクアレイ装置と、スナップショットとして提供する記憶領域を有するディスクアレイ装置をそれぞれ $nD+1P$ と $mD+1P$ のディスク構成で組むことが可能となり、各ディスクアレイ装置を互いに柔軟な構成とすることができる。

## 【0091】

又、本実施形態ではコンピュータ100、ディスクアレイ装置201及びディスクアレイ装置202を接続するインターフェースをファイバチャネルとしたが、他のインターフェース、例えばSCSI、インターネットに対応するインターフェースであってもよい。

## 【0092】

以上、本発明の第一及び第二の実施形態について説明した。

## 【0093】

尚、本発明においては、ミラー元LU、ミラー先LUのどちらか片方、もしくは、両方が、RAID5以外のRAIDレベルの構成であってもよい。この場合、 $nD+1P$ のRAID制御を行う $n$ -RAID制御サブプログラム236と、 $mD+1P$ のRAID制御を行う $m$ -RAID制御サブプログラム237が、それぞれミラー元LUとミラー先LUのRAIDレベルに合わせた制御を実施すればよい。

## 【0094】

また、ミラー先LUを複数設けた多重ミラーにおいても本発明は適用可能である。この場合、更新位置管理表222の更新ビットの欄をミラー先LU分だけ設け、LUミラーサブプログラム231がミラー元LUに対するアクセスを複数のミラー先LUに多重化する動作をすればよい。また、多重化したミラー先LUは任意の2以上の整数 $t$ を持つ $tD+1P$ のディスクアレイで構成される。そして、当該 $tD+1P$ 構成のディスクアレイをを制御することができるRAID制御手段をディスクアレイ装置に設ける。なお、ミラー先LUのうち1つ以上が、ミラー元LUの $nD+1P$ 構成と異なるRAID構成であればよい。

## 【0095】

## 【発明の効果】

以上述べたように、本発明によれば、スナップショット取得のために二重化運用されているディスクアレイにおいて、 $n$ と $m$ を異なる2以上の整数とした場合、オリジナルデータの記憶領域を有するディスクアレイと、スナップショットとして提供する記憶領域を有するディスクアレイをそれぞれ $nD+1P$ と $mD+1P$ のディスク構成で組むことが可能となり、各ディスクアレイが互いに柔軟な構成を採ることができるという効果がある。

## 【図面の簡単な説明】

## 【図1】

第1の実施形態におけるシステム構成図である。

## 【図2】

第1の実施形態における更新位置管理表の説明図である。

## 【図3】

第1の実施形態におけるスナップショット取得／削除フローである。

## 【図4】

第1の実施形態におけるミラー再同期サブプログラムの動作フローである。

## 【図5】

第1の実施形態におけるミラー元LUとミラー先LUの関係( $n < m$ )を示す図である。

## 【図6】

第1の実施形態におけるミラー元LUとミラー先LUの関係( $n > m$ )を示す図である。

## 【図7】

第2の実施形態におけるシステム構成図である。

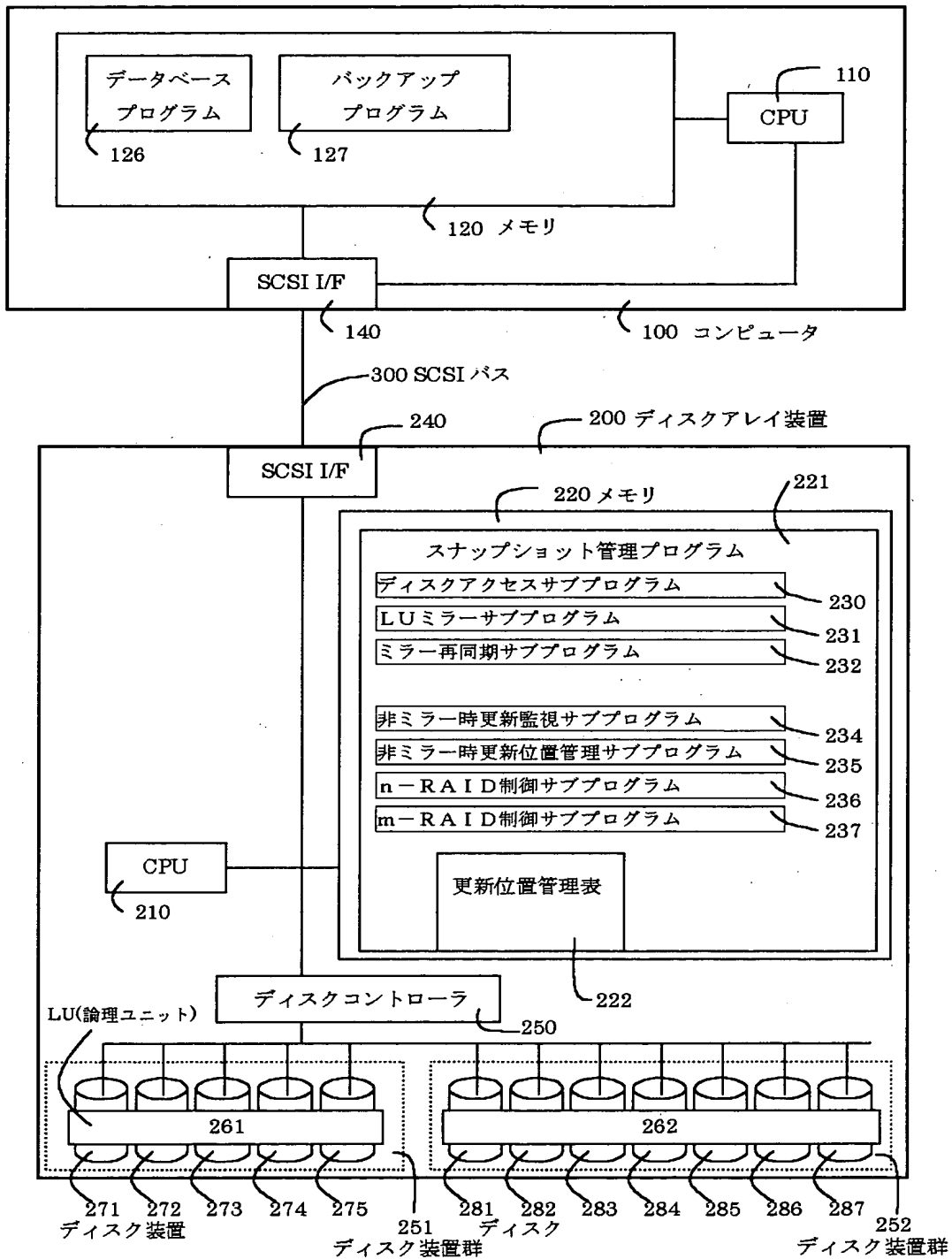
## 【符号の説明】

100…コンピュータ、200…ディスクアレイ装置、221…スナップショット管理プログラム、222…更新位置管理表、236… $n$ -RAID制御サブプログラム、237… $m$ -RAID制御サブプログラム、271～275、281～287…ディスク装置。

【書類名】 図面

【図 1】

図 1



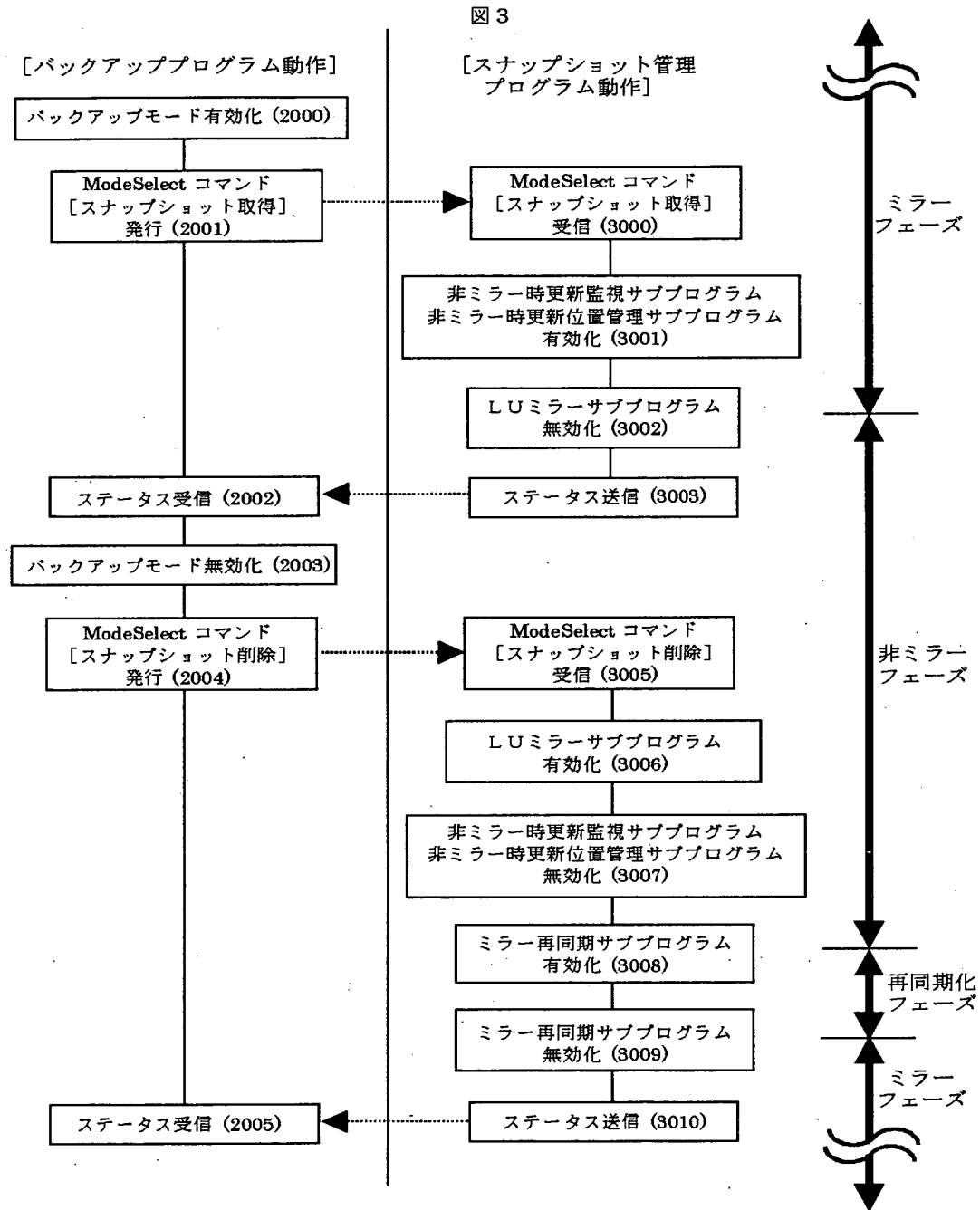
【図 2】

図 2

LBA セット番号	更新ビット
0	0
1	1
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0

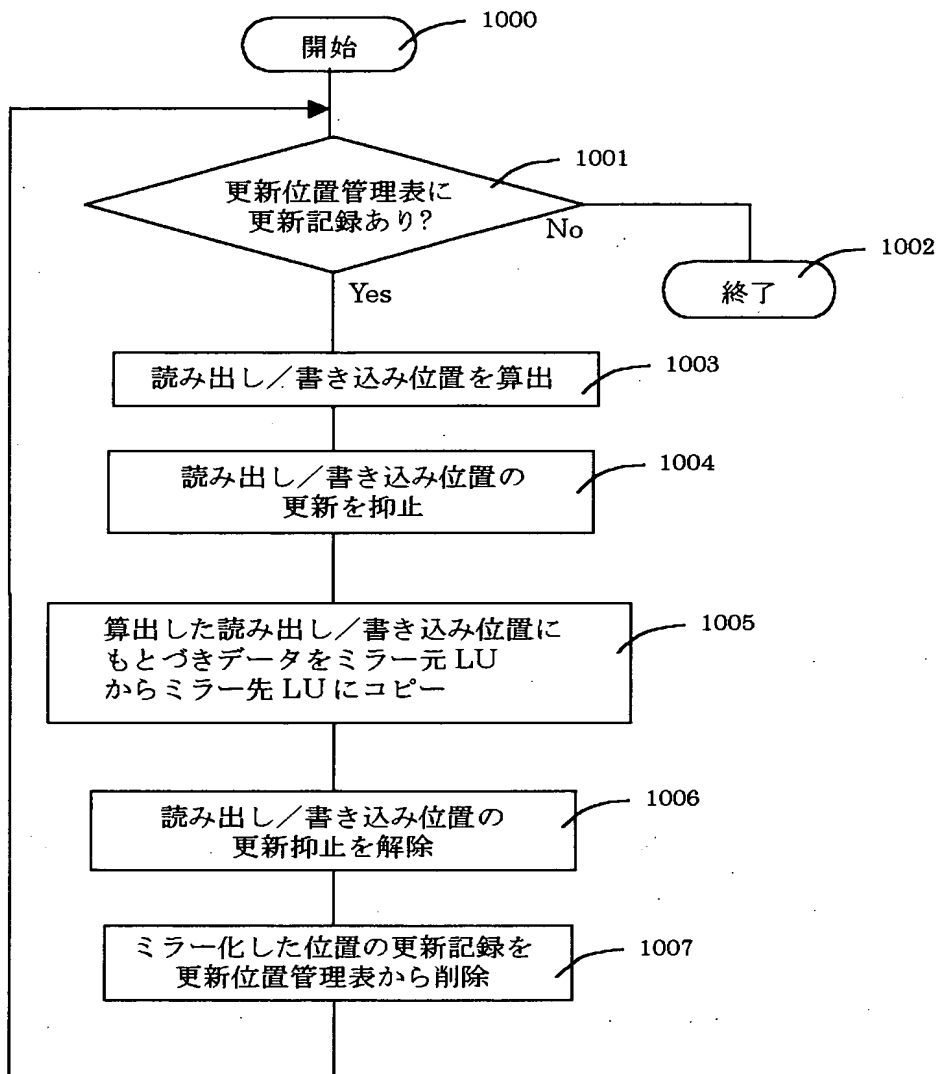


【図 3】



【図 4】

図 4



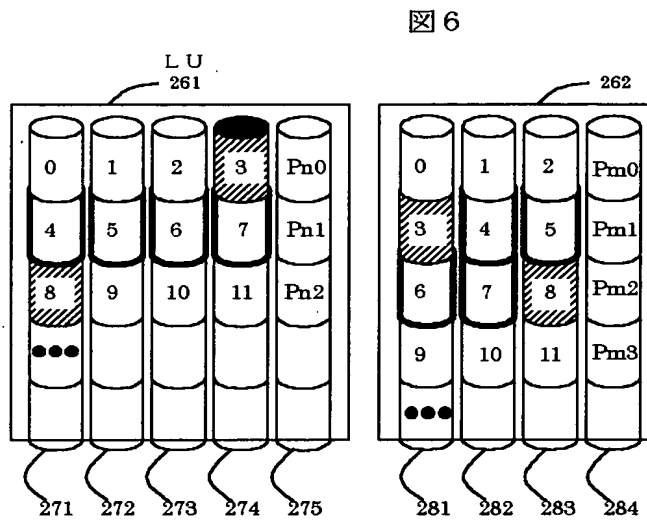
【図 5】

図 5

更新位置管理表  
222

LU261 のストライプ区切り	LU262 のストライプ区切り	LBA セット番号	更新ビット
$\uparrow n \times 0$	$\uparrow m \times 0$	0	0
		1	0
		2	0
		3	0
$\uparrow n \times 1$	$\uparrow m \times 1$	4	0
		5	0
		6	0
		7	1
$\uparrow n \times 2$	$\uparrow m \times 1$	8	1
		9	1
		10	1
		11	0
$\uparrow n \times 3$	$\uparrow m \times 2$	12	0
		13	0
		14	0
		15	0

【図 6】



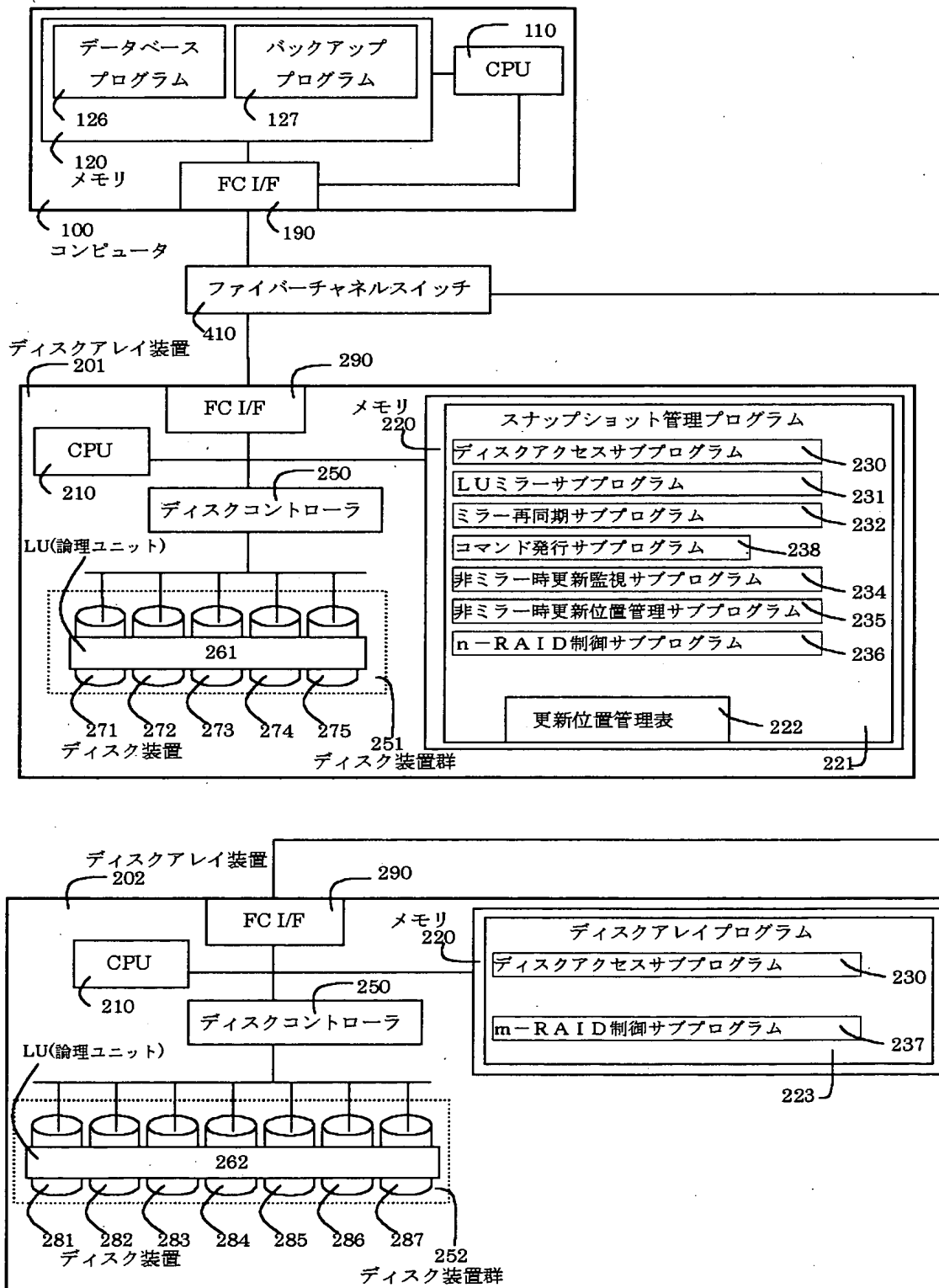
更新位置管理表

222

LU261 のストライプ区切り	LU262 のストライプ区切り	LBA セット番号	更新ビット
$\uparrow n \times 0$	$\uparrow m \times 0$	0	0
		1	0
		2	0
$\uparrow n \times 1$	$\uparrow m \times 1$	3	0
		4	1
		5	1
$\uparrow n \times 2$	$\uparrow m \times 2$	6	1
		7	1
		8	0
$\uparrow n \times 2$	$\uparrow m \times 3$	9	0
		10	0
		11	0

【図 7】

図 7



【書類名】 要約書

【要約】

【課題】

スナップショット取得のために二重化運用しているディスクアレイにおいて、オリジナルデータの記憶領域と、スナップショットとして提供する記憶領域とでは、性能、信頼性、コスト等の観点から要求されることが異なるため、性能、記録密度、容量、信頼性等が異なるディスク装置で各ディスクアレイを構成する必要がある。

【解決手段】

$nD+1P$ で構成する複数の記憶媒体上の記憶領域であるミラー元LUと、 $mD+1P$ で構成する複数の記憶媒体上の記憶領域であるミラー先LUと、 $nD+1P$ のRAID制御を行う $n$ -RAID制御サブプログラムと、 $mD+1P$ のRAID制御を行う $m$ -RAID制御サブプログラムと、コンピュータがデータ書き込みを要求したときにミラー元LUとミラー先LUに書き込んで二重化するLUミラーサブプログラムとを設ける。なお、 $m$ および $n$ は2以上の整数であって、 $m$ と $n$ は異なる値である。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所